Working Paper 95-18
Statistics and Econometrics Series 4
May 1995

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe, Madrid (Spain)
Fax (341) 624-9849

ON THE EXPLOSION BREAKDOWN RATE OF THE MAXIMUM BIAS FUNCTION OF SOME SCALE AND LOCATION ESTIMATES

José R. Berrendero, Juan Romo and Rubén Zamar [1]

Abstract

The breakdown point of an estimate, ϵ^*, gives the location at which its maxbias function, $B(\epsilon)$, explodes. The gross-error-sensitivity, GES, gives the slope of the maxbias curve at zero. Therefore, the pair (ϵ^*, GES) constitutes a useful summary of the main features of $B(\epsilon)$. We will show here that a simple and insightful summary can also be obtained by looking at the limiting relative behavior of the maxbias curve when ϵ approaches ϵ^*. We will demonstrate the usefulness of this measure and derive formulas to compute it for M-estimates of scale, dispersion and location. We will show that, likewise the gross-error-sensitivity, the breakdown rate can be straightforwardly derived from the M-estimate score function. Consequently, the maxbias behavior for large fractions of contamination can be inferred from the general shape of the M-estimate's score function. Conversely, the score function can be shaped so that some desired maxbias features are obtained.

Keywords:
Breakdown point, Bias robustness, Maxbias curve, M-estimates, Location, Scale.

[1]Berrendero and Romo, Department of Statistics and Econometrics, Universidad Carlos III de Madrid and Zamar, Department of Statistics, University of British Columbia.

1. Introduction

Suppose $X_1, \ldots, X_n$ is a sample from a distribution F in the contamination neighborhood

$$V_\epsilon(F^\theta) = \{F : F = (1-\epsilon)F^\theta + \epsilon H,\ H \text{ arbitrary distribution}\}, \quad 0 \le \epsilon \le 1. \tag{1}$$

Therefore, a proportion $1-\epsilon$ of the data comes from a distribution F^θ which belongs to a specified parametric family and the remainder come from an arbitrary distribution H. We can interpret this last proportion of data as outliers which may distort the inferences on the parameter θ.

We consider estimates T_n of θ which, under mild regularity conditions, converge a.s. [F] to the asymptotic functional $T(F)$. The functional $T(F)$ is well defined on the set of distributions for which the almost sure convergence holds and this set usually includes the family $V_\epsilon(F^\theta)$. The estimate T_n can be then represented as the functional evaluated at the empirical distribution function of the data, that is, $T_n = T(F_n)$.

Note that, due to the presence of outliers, $T(F)$ is not necessarily equal to θ for all F in $V_\epsilon(F^\theta)$. For this reason, we must consider the asymptotic bias, $b_T(F) = d(T(F), \theta)$, and its supremum,

$$B_T(\epsilon) = \sup_{F \in V_\epsilon(F^\theta)} b_T(F), \tag{2}$$

where d measures the distance between the true value of the parameter and the asymptotic value of the estimate. $B_T(\epsilon)$ represents the maximum possible perturbation of $T(F)$ when F ranges over the neighborhood and therefore it carries relevant information on the robustness properties of T. Huber (1964, 1981) introduced the concept of maximum asymptotic bias in the location model setup. Later, Martin and Zamar (1989) and Martin, Yohai and Zamar (1989) defined the maxbias curve (a plot of $B_T(\epsilon)$ versus ϵ) and computed it for several types of scale and regression estimates.

In this paper we are concerned with the issue of summarizing the information contained in the maxbias curve. There are two well-known robustness measures that summarize important aspects of this curve, both proposed by Hampel (1974): the gross error sensitivity (γ^*) and the breakdown point (ϵ^*). Although Hampel's original definition is slightly different, the gross error sensitivity turns out to be, under regularity conditions, equal to $B_T'(0)$. Therefore

$$B_T(\epsilon) \approx \gamma^* \epsilon + o(\epsilon), \text{ as } \epsilon \to 0.$$

Hence, γ^* conveys information on the behavior of $B_T(\epsilon)$ for small values of ϵ. On the other hand, the breakdown point can be defined as

$$\epsilon^* = \sup\{\epsilon : \ \Delta_T(\epsilon) > 0\}, \tag{3}$$

where $\Delta_T(\epsilon) = |B_T(1) - B_T(\epsilon)|$ if $B_T(\epsilon) < \infty$ and $\Delta_T(\epsilon) = 0$ otherwise. Since $B_T(1)$ is totally determined by the contamination distribution H which is arbitrary (see equation (1)), a proportion of contamination greater than ϵ^* may cause the estimate to take a completely arbitrary value. In this sense, ϵ^* represents the maximum proportion of contamination that the estimate can tolerate. In principle, one could expect that two estimates with similar γ^* and ϵ^* have similar maxbias curves, but this is not necessarily the case as illustrated in the following example.

Example 1. In this example we take $\theta = \sigma$ and $F^\sigma(x) = \Phi(x/\sigma)$, where Φ denotes the standard normal distribution function (see equation (1)). We consider the class of scale M–estimates (see Huber, 1964) defined as the solution, S_n, to the equation

$$\sum_{i=1}^{n} \chi(X_i/S_n) = b, \tag{4}$$

where χ is a specified function and b is equal to $E_\Phi \chi(X)$. We recall that in the context of scale M–estimates we must consider two types of maximum asymptotic biases: the explosion bias due to outliers, and the implosion bias due to inliers (see Section 2). We have considered three scale M–estimates, $\hat{\sigma}_H$, $\hat{\sigma}_L$, and $\hat{\sigma}_T$, with breakdown point 0.5 and similar explosion sensitivities (see Table 1). The corresponding score functions and tuning constants are reported in Table 1.

(Table 1 about here)

Since the three estimates have similar gross error sensitivities, their maxbias curves are similar for small values of ϵ (see Figure 1). On the other hand, the maxbias curves of $\hat{\sigma}_H$ and $\hat{\sigma}_T$ are fairly similar for all ϵ , but that of $\hat{\sigma}_L$ is quite different for values of ϵ greater than 0.25, say. Therefore, having the same breakdown point cannot be taken as an indication that the maxbias curves will have similar behaviors for large values of ϵ.

(Figure 1 about here)

Given two estimates, T_1 and T_2, with maxbias curves $B_{T_1}(\epsilon)$ and $B_{T_2}(\epsilon)$, and the same breakdown point, ϵ^*, the *relative breakdown rate* of T_1 and T_2 is defined as follows:

$$RBR(T_1, T_2) = \lim_{\epsilon \to \epsilon^*} \frac{B_{T_1}(\epsilon)}{B_{T_2}(\epsilon)}.$$

The *breakdown rate* of T_1 (with respect to T_0) is now defined as

$$BR(T_1) = RBR(T_1, T_0)$$

where T_0 is some fixed baseline estimate, for example, the minimax–bias estimate with breakdown point ϵ^*. If either $BR(T_1)$ or $BR(T_2)$ is finite, then

$$RBR(T_1, T_2) = BR(T_1)/BR(T_2)$$

and the RBR does not convey any additional information. However, there are situations where $BR(T_1)$ and $BR(T_2)$ are both equal to infinity and in this case $RBR(T_1, T_2)$ allows a direct comparison of T_1 and T_2. The BR and RBR were defined and calculated for some robust regression estimates by Mazzi (1992).

Returning to Example 1, it can be easily obtained, using the results of Corollary 1 in Section 2, that $RBR(\chi_T, \chi_H) = 1.16$ and $RBR(\chi_L, \chi_H) = \infty$. This accurately quantifies the qualitative features - analogies and discrepancies - already observed in Figure 1. Moreover, notice that these results are entirely determined by the behavior near zero of the corresponding score functions. In other words, we can "predict" the large–ϵ–maxbias behavior of an scale M–estimate by observing the shape of its score function near zero.

The rest of the paper is organized as follows. Section 2 deals with the breakdown rate for both M– and τ–estimates of scale, addressing the explosion and implosion breakdown rates separately. In Section 3 we derive the breakdown rate for several important dispersion estimates, including the SHORTH and the MAD. Section 4 considers some questions concerning the breakdown rate of M–estimates of location. Finally, Section 5 comments on some open questions and gives some final remarks.

2. Breakdown rate for some scale estimates

In this section we will derive the BR and RBR for two classes of scale estimates. First, we consider the widely used class of scale M–estimates. To improve the efficiency of these estimates preserving their robustness properties we can use the scale τ–estimates, which are considered in the second part of this section.

2.1 Breakdown rates for scale M–estimates

Consider the class of M–estimates of the scale of a positive random variable (however see Remark 1 below) with distribution function F in $V_\epsilon(F^\sigma)$, where $F^\sigma(x) = F_0(x/\sigma)$ and F_0 satisfying the following assumption:

Assumption 1.

(a) F_0 *is strictly increasing in* $(0, \infty)$ *and has a density* f_0.

(b) $\lim_{x\to\infty} f_0(x)/[1 - F_0(x)] > 0$.

Assumption 1(b) can be reworded by saying that the failure rate, $\lambda(x) = f_0(x)/[1 - F_0(x)]$, does not vanish at infinity. This condition is satisfied by many F_0 of practical importance, including the folded normal distribution, with $\lambda(x) = \varphi(x)/[1 - |\Phi(x)] \to \infty$ as $x \to \infty$ and the exponential distribution, with $\lambda(x) = 1$.

We will assume that the score function χ appearing in the definition (4) of the scale estimate satisfies the following assumption:

Assumption 2. $\chi(y)$ *is increasing for positive values of* y*, bounded with* $\chi(\infty) = 1$, $\chi(0) = 0$, *and continuous (except, perhaps, in a finite number of points).*

A scale M–estimate with score function χ satisfying Assumption 2 converges almost surely to the functional

$$S(F) = \inf\{s : \int_0^\infty \chi(y/s)dF(s) < b\},$$

where $b = E_{F_0}\chi(Y)$. In this case, all the appropriate scale–invariant measures of discrepancy between $S(F)$ and σ are functions of the ratio $S(F)/\sigma$. Since in addition the scale M–estimates are scale equivariant, we can assume without loss of generality that $\sigma = 1$.

As already mentioned in the introduction, we must consider two kinds of maximum asymptotic biases: the overestimation bias, $B_\chi^+(\epsilon)$, caused by outliers, and the underestimation bias, $B_\chi^-(\epsilon)$, caused by inliers. The corresponding breakdown points are the explosion breakdown point, $\epsilon^+ = \sup\{\epsilon : B_\chi^+(\epsilon) < \infty\}$, and the implosion breakdown point, $\epsilon^- = \sup\{\epsilon : B_\chi^-(\epsilon) > 0\}$. Martin and Zamar (1989) found that

$$B_\chi^+(\epsilon) = \sup_{F\in V_\epsilon} S(F) = h_\chi^{-1}[(b-\epsilon)/(1-\epsilon)], \tag{5}$$

and

$$B_\chi^-(\epsilon) = \inf_{F\in V_\epsilon} S(F) = h_\chi^{-1}[b/(1-\epsilon)], \tag{6}$$

where $h_\chi(s) = \int_0^\infty \chi(y/s)dF_0(s)$. Therefore, we will consider two kinds of breakdown rates, $BR^+(\chi)$ and $BR^-(\chi)$, corresponding to explosion and implosion biases, respectively. Since $\lim_{s\to\infty} h_\chi(s) = 0$ and $\lim_{s\to 0} h_\chi(s) = 1$, it follows that $\epsilon^+ = b$ and $\epsilon^- = 1 - b$. Therefore, the breakdown point of the estimate (see equation (3)) is equal to $\epsilon^* = \min\{\epsilon^+, \epsilon^-\} = \min\{b, 1-b\}$. This last expression agrees with the formula derived earlier by Huber (1981).

Consider the class C_b of χ functions such that $b = E_{F_0}\chi(Y)$. All functions in C_b have the same explosion and implosion breakdown points and for each $0 < b < 1$, the jump function

$$\chi_a(y) = \begin{cases} 0 & \text{if } |y| < a \\ 1 & \text{if } |y| \geq a \end{cases},$$

with $a = F_0^{-1}(1-b)$, belongs to C_b. Martin and Zamar (1989) showed that, if $f_0(sx)/f_0(x)$ is decreasing in x for $s > 1$, then

$$B_\chi^+(\epsilon) \geq B_a^+(\epsilon) \quad \text{for all} \quad 0 < \epsilon < b, \text{ and}$$

$$B_\chi^-(\epsilon) \leq B_a^-(\epsilon) \quad \text{for all} \quad 0 < \epsilon < 1-b,$$

that is, regarding explosion and implosion maxbias curves, χ_a dominates all the the score functions in C_b. In view of this property, it seems convenient to choose χ_a as the baseline in the definition of $BR^+(\chi)$ and $BR^-(\chi)$, for $\chi \in C_b$.

REMARK 1. All the results obtained in this section remain valid if we replace Assumption 1(a) by the following: F_0 is strictly increasing and has a symmetric and unimodal density f_0. So, we can apply these results to many random variables that take values over the entire real line as the normal or the double exponential. Only three minor adjustments are needed: the function χ must be assumed to be even, the baseline jump function must jump at $\pm F_0^{-1}(1-b/2)$ and the constant c_k of Theorem 1 below must be multiplied by 2.

2.1.1 Explosion breakdown rates

Most χ functions proposed in the literature –in fact all the χ functions that we are aware of– can be written in the form

$$\chi(y) = ry^k + o(y^k), \qquad \text{as} \quad y \to 0^+$$

for some integer k and $r > 0$. Therefore, $\chi^{(j)}(0) = 0$ for all $j < k$ and $\chi^{(k)}(0) = r > 0$. In such case we will say that the *local order of* χ *(at zero)* is k. The χ functions of the jump type are a limit case because $\chi^{(k)}(0) = 0$ for all k. In this case we say that the local order is $k = \infty$. The reader can easily verify that χ_L (see Example 1) has local order $k = 1$ and χ_H and χ_T have local order $k = 2$.

Example 2. We now consider two score functions from Croux (1994), $\chi_C(y) = y^2/(y^2+c^2)$ and $\chi_W(y) = 1 - \exp(-y^2/c)$, which unlike χ_L, χ_H and χ_T are not a polynomial near zero. These functions are also to be used for illustrations later on. Since $\chi_C'(0^+) = \chi_W'(0^+) = 0$, $\chi_C^{(2)}(0^+) = 2/c^2$, and $\chi_W^{(2)}(0^+) = 2/c$, the local order of χ_C and χ_W is $k = 2$. Observe that χ_C and χ_W are strictly monotone on $[0,\infty)$ and reach their maximum value one at infinity. The corresponding scale estimates $\hat{\sigma}_C$ and $\hat{\sigma}_W$ have breakdown point 0.5 if $c = 0.61$ and $c = 0.66$, respectively.

In the next theorem we give an approximation (valid for large ϵ) of the explosion maxbias curves for a wide class of scale M–estimates. From this approximation we will easily compute the BR and RBR of these estimates.

Theorem 1. *Suppose that F_0 satisfies Assumption 1 and that, for some $R > 0$, $f_0(x) = O(e^{-rx})$ for all $0 < r < R$. Suppose also that $\chi \in C_b$ satisfies Assumption 2. If the local order of χ is $k < \infty$, then*

$$\lim_{\epsilon \to b} \left(\frac{1-\epsilon}{b-\epsilon}\right)^{-1/k} B^+_\chi(\epsilon) = c_k(\chi), \tag{7}$$

where c_k is a constant independent of ϵ given by

$$c_k = \left[\frac{\chi^{(k)}(0^+)}{k!}\int_0^\infty y^k f_0(y)dy\right]^{1/k}. \tag{8}$$

The cases $k = 1$ and $k = 2$ are particularly important because many proposed χ functions have these orders. If the local order of χ is $k = 1$, then $B^+_\chi(\epsilon) \propto (b-\epsilon)^{-1}$, for large ϵ, and if $k = 2$ then $B^+_\chi(\epsilon) \propto (b-\epsilon)^{-1/2}$, for large ϵ. In general, score functions with large local order produce estimates with good maxbias performance for large ϵ. Since functions with large local orders are flat near zero, it follows that the flatness of χ in a neighborhood of zero yields a better behavior of the explosion bias for large fractions of contamination.

REMARK 2. A question of perhaps some theoretical interest is what happens when the local order of χ is not an integer. For example, we could consider a function whose local behavior near zero is approximately $y^{1/2}$. Indeed, when $\chi(y) = ry^\alpha + o(y^\alpha)$ in a neighborhood of zero, for some $\alpha > 0$, it can be proved, following the lines of the proof of Theorem 1, that

$$\lim_{\epsilon \to b} \left(\frac{1-\epsilon}{b-\epsilon}\right)^{-1/\alpha} B^+_\chi(\epsilon) = c_\alpha(\chi),$$

where $c_\alpha = \left(r\int_0^\infty y^\alpha f_0(y)dy\right)^{1/\alpha}$.

Corollary 1. *Suppose that the local order of χ_i is $k_i < \infty$ ($i = 1,2$) and that the assumptions of Theorem 1 hold. Then* *(a) If* $k_1 = k_2 = k$, $RBR^+(\chi_1,\chi_2) = [\chi_1^{(k)}(0^+)/\chi_2^{(k)}(0^+)]^{1/k}$.

(b) If $k_1 < k_2$, $RBR^+(\chi_1,\chi_2) = \infty$.

Example 1. *(Continued)* The approximations given in Theorem 1 corresponding to the estimates $\hat{\sigma}_L$, $\hat{\sigma}_T$, and $\hat{\sigma}_H$ are displayed in Figure 2 for $\epsilon > 0.2$.

(Figure 2 about here)

By Theorem 1, the explosion bias of $\hat{\sigma}_L$ behaves (except for constants and for large ϵ) as the square of the explosion biases of $\hat{\sigma}_T$ and $\hat{\sigma}_H$. More precisely, when F_0 is normal, it follows that

$$\begin{aligned}\lim_{\epsilon\to 0.5}[B^+_{\chi_H}(\epsilon)]^2/B^+_{\chi_L}(\epsilon) &= 1.35, \text{ and}\\ \lim_{\epsilon\to 0.5}[B^+_{\chi_T}(\epsilon)]^2/B^+_{\chi_L}(\epsilon) &= 1.83.\end{aligned}$$

In Table 2 we compare seven scale M–estimates with breakdown point equal to 0.5. The score functions χ_L, χ_T, χ_H, χ_W and χ_C were defined before (see Table 1 Example 2). The new score functions, $\chi_4(y) = \min(y^4/c^4, 1)$, $c = 0.85$, and

$$\chi_{CR}(y) = \begin{cases} 1, & \text{if } |y| \geq k\\ by^2 + c, & \text{if } \delta \leq |y| < k\\ a|y|, & \text{if } 0 \leq |y| < \delta \end{cases},$$

are included to illustrate the relationship between efficiency and breakdown rate. The function χ_{CR} was initially introduced by Croux (1994) to show that there is not trade–off between high breakdown point and efficiency in the class of scale M–estimates.

(Table 2 about here)

By choosing the constants a, b, c, k and δ appropriately (see details in Croux, 1994) this estimate has breakdown point equal to 0.5 and an arbitrarily high efficiency. Croux also pointed out that the actual trade-off is between maximum bias and efficiency. Since Croux's estimate has local order $k = 1$ and χ_H (the one taken as a baseline in Table 2) has local order $k = 2$, by Corollary 1 (b), $RBR^+(\chi_{CR},\chi_H) = \infty$. Therefore, the trade-off between efficiency and robustness is also reflected by the explosion breakdown rate. From the explosion breakdown rate point of view, χ_{CR}

is only comparable to χ_L among the estimates considered in Table 2 (both have local order equal to 1). Using Corollary 1 we obtain $RBR^+(\chi_{CR}, \chi_L) = 6.18$. For the more efficient Croux's estimates with efficiencies 92% and 94% the relative breakdown rates are $RBR^+(\chi_{CR}, \chi_L) = 60.2$ and $RBR^+(\chi_{CR}, \chi_L) = 602.7$, respectively. To complete our discussion of breakdown rate and efficiency observe that, since $\chi_4(y)$ has local order $k = 4$, its explosion breakdown rate is very good but its efficiency is very low. Our conjecture is that low explosion breakdown rate and high efficiency cannot be simultaneously achieved in the class of scale M–estimates. Another immediate consequence of Theorem 1 is the following result.

Corollary 2. *Suppose that the assumptions of Theorem 1 are satisfied. If the local order of χ at zero is $k < \infty$, then it holds that $BR^+(\chi) = \infty$.*

If the χ function is totally flat on some neighborhood of zero then the local order of χ is equal to ∞ and the condition of Corollary 2 does not hold ($\chi^{(j)}(0) = 0$ for all j). We will show that in this case the BR is finite and given by the formula $BR^+(\chi) = a/m$, where $a = F_0^{-1}(1-b)$ and $m = \sup\{y : \chi(y) = 0\}$. To see that $BR^+(\chi) \le a/m$, just notice that

$$\begin{aligned} h_\chi(s) &= \int_0^\infty \chi(y/s)dF_0(y) = \int_{ms}^\infty \chi(y/s)f_0(y)dy \\ &\le \int_{ms}^\infty f_0(y)dy = 1 - F_0(ms). \end{aligned}$$

Replacing s by $B_\chi^+(\epsilon)$ and rearranging terms, we have

$$B_\chi^+(\epsilon) \le \frac{1}{m}F_0^{-1}\left(1 - \frac{b-\epsilon}{1-\epsilon}\right) = \frac{1}{m}F_0^{-1}\left(\frac{1-b}{1-\epsilon}\right).$$

The desired inequality follows now because, by (5),

$$B_{\chi_a}^+(\epsilon) = \frac{1}{a}F_0^{-1}\left(1 - \frac{b-\epsilon}{1-\epsilon}\right) = \frac{1}{a}F_0^{-1}\left(\frac{1-b}{1-\epsilon}\right).$$

The converse inequality is proved in the Appendix. Therefore, we have proved the following theorem.

Theorem 2. *Suppose that F_0 satisfies Assumption 1, F_0 has a decreasing density f_0 and $\chi \in C_b$ satisfies Assumption 2. If χ is constant on a neighborhood of zero then $BR^+(\chi) = a/m$, where $a = F_0^{-1}(1-b)$ and $m = \sup\{y : \chi(y) = 0\}$.*

As a consequence, if we want to obtain estimates with explosion bias comparable to that of the minimax bias estimate for large ϵ, we must restrict attention to χ functions which are flat near zero. Notice that since $\chi \in C_b$, then $m < a$ and it follows that $BR^+(\chi) > 1$.

2.1.2 Implosion breakdown rates

The next theorem gives the implosion breakdown rate for scale estimates with score function ultimately constant. Almost all the χ functions proposed in the robustness literature satisfy this condition.

Theorem 3. *Suppose that F_0 satisfies Assumption 1(a) and has a decreasing density f_0. Suppose also that $\chi \in C_b$ satisfies Assumption 2. If there exists $A > 0$ such that $\chi(y) = 1$ for every $y \geq A$, then*

$$BR^{-}(\chi) = a\left[\int_0^A [1 - \chi(y)]dy\right]^{-1}, \qquad \text{where} \quad a = F_0^{-1}(1 - b).$$

Theorem 3 does not apply to functions that reach their maximum only at infinity. In the following theorem we give the implosion breakdown rate for differentiable χ functions that are not ultimately constant but whose derivative decreases fast enough.

Theorem 4. *Suppose that F_0 satisfies Assumption 1(a), $f_0(0) > 0$ and $\chi \in C_b$ satisfies Assumption 2. If χ is differentiable, $\lim_{s\to 0} \chi'(y/s)s^{-2} = 0$, and $\int_0^\infty \chi'(y)ydy < \infty$, then*

$$BR^{-}(\chi) = a\left[\int_0^\infty \chi'(y)ydy\right]^{-1}, \qquad \text{where} \quad a = F_0^{-1}(1 - b).$$

Notice that the functions χ_C and χ_W from Table 2 are not ultimately constant but satisfy the assumptions of Theorem 4. Also notice that if the function χ satisfies the assumptions of both Theorems 3 and 4, then the results of these two theorems are identical:

$$\int_0^A \chi'(y)ydy = A - \int_0^A \chi(y)dy = \int_0^A [1 - \chi(y)]dy,$$

2.2 Breakdown rates for scale τ-estimates

The class of scale τ-estimates combines the properties of high efficiency and breakdown point. These estimates were defined by Yohai and Zamar (1988) in the context of robust regression. Let $\chi_i \in C_b$, for $i = 1, 2$ be a pair of score functions. The τ-scale of F is defined as

$$\tau(F) = S(F)\left[\frac{1}{b_2}E_F\chi_2\left(\frac{Y}{S(F)}\right)\right]^{1/2},$$

where $S(F)$ is the scale M–estimate based on χ_1, that is, $S(F)$ is the solution of

$$E_F\chi_1\left(\frac{Y}{S(F)}\right) = b_1.$$

To compute the τ–estimate of scale, just replace F by the empirical distribution function F_n in the definition above.

The breakdown point of a τ–estimate is $\epsilon^* = \min\{b_1, 1 - b_1\}$. That is, a τ–estimate inherits the breakdown point of the scale M–estimate based on χ_1. Rousseeuw and Croux (1994) have found the maxbias curves of τ–estimates with respect to outliers and inliers when the function $s \to s^2 E_{F_0}\chi_2(Y/s)$ is increasing for $s \geq 0$. In that case, a direct application of their formulas (3.9) and (3.10) yields

$$RBR^+(\tau, S) = RBR^-(\tau, S) = (b_1/b_2)^{1/2}.$$

Therefore, the τ–estimates do not inherit the breakdown rate of the initial scale M–estimate, S(F). In fact, since b_2 is selected to obtain efficiency whereas b_1 is chosen to obtain high breakdown point, in general, $b_1/b_2 > 1$.

For example, if χ_1 and χ_2 are given by χ_T (see Table 1) with the tuning constants fitted to get $b_1 = 0.5$ and $b_2 = 0.16$ (to obtain 95% efficiency under the normal model), then $RBR^+(\tau, S) = RBR^-(\tau, S) = 1.76$.

3. Breakdown rates for dispersion estimates

We now consider the location-dispersion model with central distribution

$$F^{\mu,\sigma}(x) = F_0[(x-\mu)/\sigma]$$

where F_0 is a known distribution function that satisfies the following assumption:

Assumption 3. *F_0 is strictly increasing, has a symmetric and unimodal density f_0 and satisfies Assumption 1(b).*

In this section we will derive the breakdown rate for dispersion estimates of several types. The dispersion estimates can be viewed as scale estimates applied to the centered data, $X_1 - T_n, \ldots, X_n - T_n$, where the location estimate, T_n, can be defined in different ways. One way is to choose T_n in advance in some ad hoc way (e.g. T_n = Median). A well known example of an estimate of this type is the MAD,

$$M_n = c\ \mathrm{med}_i\{|x_i - \mathrm{med}_j x_j|\}, \qquad c = 1/F_0^{-1}(0.75),$$

where T_n = Median, the score function is of jump type and $b = 0.5$. A second way is to calculate T_n simultaneously with the dispersion estimate as in Huber's Proposal 2 (see Huber, 1964). Since simultaneous estimates of location and dispersion have poor robustness properties we shall not consider these type of estimates here. A third way is to define T_n as the value of t that minimizes the M–scale, $D_n(F,t)$, of the shifted data, $X_1 - t, \ldots, X_n - t$; that is,

$$T_n = \operatorname{argmin}_{t \in \Re} D_n(F,t).$$

These location estimates are called location S–estimates and were defined (in the regression setup) by Rousseeuw and Yohai (1984); the corresponding dispersion estimates, $D_n(F, T_n)$, are called S–estimates of dispersion. When the score function is of the jump type with $b = 0.5$, $D_n(F, T_n) =$ SHORTH (see Rousseeuw and Leroy, 1987).

The definitions of explosion and implosion maxbias curves for dispersion estimates are analogous to those given in Section 2 for scale estimates. Moreover, the explosion and implosion maxbias curves of the dispersion S–estimates are equal to the maxbias curves of the corresponding scale M–estimates (see Martin and Zamar, 1993b). Therefore, all the results obtained in Section 2 for scale M–estimates also hold for dispersion S–estimates.

Martin and Zamar (1993b) also showed that the implosion maxbias curves of the dispersion S-estimates and the corresponding dispersion M–estimates with ad hoc centering and the same score function are equal. Therefore, the implosion maxbias curves for the MAD and the SHORTH are equal and $RBR^{-}(MAD, SHORTH) = 1$. Their relative explosion breakdown rate is given in the following theorem.

Theorem 5. *If F_0 satisfies Assumption 3 then $RBR^{+}(MAD, SHORTH) = 2$.*

Rousseeuw and Croux (1993) proposed the following two dispersion estimates which are more efficient under the Gaussian model than the SHORTH and the MAD and have breakdown point equal to $1/2$:

$$\begin{aligned} S_n &= k \operatorname{med}_i\{\operatorname{med}_j |X_i - X_j|\}, \text{ and} \\ Q_n &= d\ \{|X_i - X_j| : i < j\}_{(p)}, \end{aligned}$$

where $p = \binom{h}{2}$ with $h = [n/2] + 1$. The following theorem gives the breakdown rates of these estimates when F_0 is normal. These, as all the relative breakdown rates calculated in this section, are with respect to the SHORTH.

Theorem 6. *If $F_0 = \Phi$ and $c = 1/\Phi^{-1}(0.75)$, then*

(a) $BR^+(S) = 2k/c$ and $BR^+(Q) = 2^{1/2}d/c$.

(b) $BR^-(S) = k/c$ and $BR^-(Q) = [c(1 + 2^{-3/2}d^{-1})]^{-1}$.

REMARK 3. The assumption of normality is not needed to prove part (a) of Theorem 5 nor to derive $RB^-(S)$.

Table 3 contains the resulting breakdown rates when we choose the suitable constant values to obtain Fisher-consistency under the Gaussian model. Although Q_n is the most efficient, it is nót robust against inliers. On the other hand, S_n seems a reasonable compromise between good maxbias behavior and efficiency.

(Table 3 about here)

4. Breakdown rate for location M–estimates

Assume the same central model as in Section 3. Our interest is focussed now on the study of the breakdown rates for M–estimates of location. This class of estimates is defined (see Huber, 1964) as

$$T_n = \inf\{t : \sum_{i=1}^{n} \psi[(X_i - t)/D_n] < 0\},$$

where D_n is a robust scale equivariant estimate of the dispersion parameter, σ, and ψ satisfies

Assumption 4. *ψ is odd, non decreasing, bounded with $\psi(\infty) = 1$, and continuous (except, perhaps, in a finite number of points).*

Including the dispersion estimate, D_n, is needed to ensure that the location M–estimate is translation and scale equivariant. It can be shown that T_n converges almost surely to

$$T_\psi(F) = \inf\{t : \int \psi[(y - t)/D(F)]dF(y) < 0\},$$

where $D(F)$ is the asymptotic value of D_n. In this context, the appropriate measure of discrepancy between $T(F)$ and μ is the absolute value of the difference standardized by the dispersion. Then, the maximum asymptotic bias, as defined in equation (2), reduces to

$$B_\psi(\epsilon) = \sup_{F \in V_\epsilon} |(T(F) - \mu)/\sigma|.$$

Observe that, in the location case, $B(1) = \infty$ and then the breakdown point, as defined in equation (3), is

$$\epsilon^* = \sup\{\epsilon : B_T(\epsilon) < \infty\}.$$

Given that T_ψ is a location and scale equivariant estimate, we can assume without loss of generality that $\mu = 0$ and $\sigma = 1$.

Martin and Zamar (1993a) give the following expression for the bias function of T_ψ:

$$B_\psi(\epsilon) = g_\psi^{-1}[\epsilon/(1-\epsilon), D^+(\epsilon)], \tag{9}$$

where $g_\psi(t,s) = -\int_0^\infty \psi((y-t)/s)dF_0(y)$, $g_\psi^{-1}(x,s)$ is the inverse function of g_ψ with respect to its first argument and $D^+(\epsilon) = \sup_{F\in V_\epsilon} D(F)$. This expression is valid provided that $0 < D(F) \leq D^+(\epsilon) < \infty$ for every $F \in V_\epsilon$. From (9), it follows that the breakdown point of T_ψ is equal to 0.5. In particular, taking $\psi_0(y) = \mathrm{sgn}(y)$ leads to the median. We can apply (9) to obtain that the maximum bias function of the median is given by

$$B_0(\epsilon) = F_0^{-1}[0.5/(1-\epsilon)]$$

which agrees with the expression found by Huber (1964). We also recall that Huber (1964) showed that for every location and scale equivariant estimate, T,

$$B_0(\epsilon) \leq B_T(\epsilon) \quad \text{for all} \quad \epsilon \in (0, 0.5).$$

That is, the median is minimax-bias among all the location and scale equivariant estimates. Therefore, it is natural to take the median as baseline estimate in the definition of breakdown rate for location estimates.

The following theorem is the main result of this section. It gives bounds for the breakdown rate of the location M–estimates when the dispersion is estimated by an arbitrary previous estimate D with explosion breakdown rate $BR^+(D)$ (see Section 3).

Theorem 7. *Let $c = 1/F_0^{-1}(0.75)$ and suppose that Assumptions 3 and 4 hold. Then*

(a) If there exists $A > 0$ such that $\psi(y) = 1$ for every $y \geq A$,

$$\max\{1, cBR^+(D)A\} \leq BR(\psi) \leq 1 + cBR^+(D)A.$$

(b) If $\psi(y) < 1$ for every $y \in \Re$,

$$BR(\psi) = \infty.$$

Remark 4. In Lemma 3 (see Appendix) we give an example of an M–estimate whose breakdown rate is exactly equal to $1 + cBR^{+}(D)A$. Therefore, it is not possible to obtain a sharper general upper bound.

The M–estimates of location with monotone score functions can be viewed as the solutions to the problem of minimizing a convex loss function. An ultimately constant score function ψ is related to an ultimately linear loss function. In this case, as shown in the first part of the theorem, the BR increases roughly linearly with the value A beyond which the loss function is linear. The slope of this linear growth is related to the explosion breakdown rate of the previously chosen dispersion estimate. In Figure 3, lower and upper bounds are displayed as functions of A when we take the MAD or the SHORTH as previous estimates.

(Figure 3 about here)

On the other hand, a strictly increasing score function ψ is linked with a loss function that always increases more slowly than a linear function. In this case, the breakdown rate is equal to infinity as shown in part (b) of the theorem.

Hence, we have established two classes of convex loss functions with a sharply different behavior with respect to the breakdown rate. As we have seen, a good behavior for the asymptotic bias curve (for large ϵ) requires a loss function ultimately linear. An important example of estimates in this group is provided by the well-known minimax-variance Huber estimates. In fact, all the location M–estimates with some robustness optimality property that we are aware of belong to this group. Therefore, our result here simply confirms the good robustness properties of this class of M–estimates.

In Table 4, we present the middle point between the lower and the upper bounds for several common values of A and for the MAD and the SHORTH as previous dispersion estimates in the Gaussian case.

(Table 4 about here)

Notice that for $A > F_0^{-1}(0.75)$, our bounds imply that the breakdown rate using the SHORTH as dispersion estimate will always be smaller than the breakdown rate using the MAD.

Remark 5. In the less realistic case when σ is known, it is possible to obtain a somewhat surprising result following the lines of the proof of Theorem 1, part (a). When the dispersion is known and the score function is ultimately constant, the breakdown rate is equal to 1 for all values of A. Indeed, it is possible to show that the difference between the maximum biases of one of these estimates and the median is always less than or equal to A. That is, the breakdown rate deficiency of the M–estimates of location is mainly due to the estimation of the unknown dispersion parameter.

5. Concluding remarks

We want to finish this paper by pointing out some questions related to the breakdown rate that may be object of further research. First, observe that several problems in which one tries to optimize the efficiency subject to a number of robustness constraints have been considered in the statistical literature. See, for example, Hampel (1968), Hössjer (1992) and Yohai and Zamar (1992). In particular, as it is well-known, Hampel solved the following problem:

$$\min_{T \in \mathcal{T}} AV(T, F_0) \quad \text{subject to} \quad \gamma_T^* \leq \gamma_0,$$

where $\mathcal{T}$ is the class of location or scale M–estimates. That is, Hampel found the most asymptotically efficient estimate under the model among all the estimates in $\mathcal{T}$ whose sensitivity is less than or equal to γ_0.

When $\mathcal{T}$ is the class of scale M–estimates, we may have two kinds of different solutions depending on the value γ_0. If γ_0 is large, the optimal function is of the Huber type (see function χ_H in Section 2). On the other hand, for small values of γ_0 the solution turns out to be a truncated from below χ_H function and, therefore, is totally flat near zero. Applying Theorem 2, this means that if we impose a strong restriction on the sensitivity in Hampel's problem, the maximum bias curve of the optimal estimate will also have a good behavior for large values of ϵ.

However, it would be of interest to include the breakdown rate in the set of robustness restrictions, allowing at the same time larger values for the sensitivity. Regarding this point, one realizes that a problem in which the efficiency is optimized subject to restrictions on the entire maximum bias curve would be highly desirable. Unfortunately, this problem will never be solved as such because it would entail an infinite number of side constraints. On the other hand, requiring that the sensitivity and the breakdown rate are below certain bounds would, to a great extend, limit the class of allowable maximum bias curves. In this sense, an efficiency optimization problem with side constraints on the sensitivity and the breakdown rate would be a reasonably good approximation to the global optimization problem mentioned before. This is a problem that, in our opinion, deserves further attention.

Appendix. Proofs

Proof of Theorem 1: Since k is the local order of χ at zero, the following Taylor expansion is valid for $y < \delta s$.

$$\chi\left(\frac{y}{s}\right) = \chi^{(k)}(0^+)\frac{y^k}{k!s^k} + o\left(\frac{y^k}{s^k}\right).$$

Therefore, we can write the function h_χ as

$$\begin{aligned} h_\chi(s) &= \int_0^{\delta s} \chi\left(\frac{y}{s}\right) f_0(y)dy + \int_{\delta s}^{\infty} \chi\left(\frac{y}{s}\right) f_0(y)dy \\ &= \frac{\chi^{(k)}(0^+)}{k!s^k} \int_0^{\delta s} y^k f_0(y)dy + \int_0^{\delta s} o\left(\frac{y^k}{s^k}\right) f_0(y)dy \\ &+ \int_{\delta s}^{\infty} \chi\left(\frac{y}{s}\right) f_0(y)dy. \end{aligned}$$

If we multiply both sides of this equation by s^k and take limits as $s \to \infty$, we obtain

$$\lim_{s\to\infty} s^k h_\chi(s) = \frac{\chi^{(k)}(0^+)}{k!} \int_0^{\infty} y^k f_0(y)dy \tag{10}$$

since

$$\lim_{s\to\infty} \int_0^{\delta s} \frac{s^k}{y^k} o\left(\frac{y^k}{s^k}\right) y^k f_0(y)dy = 0,$$

by Dominated Convergence Theorem, and

$$\int_{\delta s}^{\infty} s^k \chi\left(\frac{y}{s}\right) f_0(y)dy \leq s^k[1 - F_0(\delta s)],$$

where due to Assumption 1(b) and the behavior of the tails of f_0, this latter expression goes to zero as $s \to \infty$. Notice also that our assumptions imply that F_0 has finite moments of all orders.

By replacing now s with $B_\chi^+(\epsilon)$ in equation (10), it follows that

$$\lim_{\epsilon\to b}[B_\chi^+(\epsilon)]^k \frac{b-\epsilon}{1-\epsilon} = \frac{\chi^{(k)}(0^+)}{k!} \int_0^{\infty} y^k f_0(y)dy = [c_k(\chi)]^k.$$

The result follows immediately from this equation.

Proof of Corollary 2: For the baseline M–estimate, we have that $s^k h_{\chi_a}(s) = s^k[1 - F_0(as)]$. Hence, taking $s = B_{\chi_a}^+(\epsilon)$ it follows that

$$\lim_{\epsilon\to b}[B_{\chi_a}^+(\epsilon)]^k \left(\frac{b-\epsilon}{1-\epsilon}\right) = 0, \text{ for every integer } k.$$

The result follows from this fact and Theorem 1.

Proof of Theorem 2: It remains to show that $BR^{+}(\chi) \geq a/m$. For $d > m$, define

$$\chi_d(y) = \begin{cases} 0 & \text{if } |y| < d \\ 1 & \text{if } |y| \geq d \end{cases}.$$

We have that

$$\begin{aligned} h_\chi(s) - h_{\chi_d}(s) &= s\int_m^\infty \chi(y) f_0(ys)dy - s\int_d^\infty f_0(ys)dy \\ &= s\int_m^d \chi(y) f_0(ys)dy - s\int_d^\infty [1-\chi(y)] f_0(ys)dy \\ &\geq s f_0(ds)\int_m^d \chi(y)dy - [1-\chi(d)][1-F_0(ds)] \\ &= [1-F_0(ds)]\left[\frac{s f_0(ds)}{1-F_0(ds)}\int_m^d \chi(y)dy - [1-\chi(d)]\right]. \end{aligned}$$

This expression is strictly positive for large enough s since, from Assumption 1(b), $\lim_{x\to\infty} x f_0(x)[1-F_0(x)]^{-1} = \infty$. Therefore, given $d > m$ we have shown that $h_\chi(s) - h_{\chi_d}(s) > 0$ for large enough s. Replacing s by $B_\chi^{+}(\epsilon)$ in this inequality and using that $h_{\chi_d}(s) = 1 - F_0(ds)$, it follows that

$$B_\chi^{+}(\epsilon) > \frac{1}{d} F_0^{-1}\left(1 - \frac{b-\epsilon}{1-\epsilon}\right),$$

for ϵ large enough and for all $d > m$. This implies $BR^{+}(\chi) \geq a/d$ for all $d > m$. Now, if $d \to m$, we deduce $BR^{+}(\chi) \geq a/m$.

Proof of Theorem 3: Let $D = \int_0^A [1-\chi(y)]dy = A - \int_0^A \chi(y)dy$. Observe that

$$\int_0^D \chi(y)dy - \int_D^A [1-\chi(y)]dy = \int_0^A \chi(y)dy - (A - D) = 0. \tag{11}$$

Define the jump function

$$\chi_D(y) = \begin{cases} 0 & \text{if } |y| < D \\ 1 & \text{if } |y| \geq D \end{cases}.$$

Notice that

$$\begin{aligned} h_\chi(s) - h_D(s) &= s\left[\int_0^D \chi(y) f_0(sy)dy - \int_D^A [1-\chi(y)] f_0(sy)dy\right] \\ &\geq s f_0(sD)\left[\int_0^D \chi(y)dy - \int_D^A [1-\chi(y)]dy\right] = 0, \end{aligned}$$

since f_0 is decreasing and applying (11). Taking $s = B_\chi^-(\epsilon)$, we obtain

$$B_\chi^-(\epsilon) > \frac{1}{D} F_0^{-1}[1 - b/(1-\epsilon)]$$

and, therefore, $BR^-(\chi) \geq a/D$.

To show the opposite inequality, observe that due to (11), for every $d < D$,

$$\int_0^d \chi(y)dy < \int_0^D \chi(y)dy = \int_D^A [1-\chi(y)]dy < \int_d^A [1-\chi(y)]dy. \tag{12}$$

Define, for $d < D$, the jump function

$$\chi_d(y) = \begin{cases} 0 & \text{if } |y| < d \\ 1 & \text{if } |y| \geq d \end{cases}.$$

We have that

$$\begin{aligned} h_\chi(s) - h_d(s) &= s\left[\int_0^d \chi(y) f_0(sy)dy - \int_d^A [1-\chi(y)] f_0(sy)dy\right] \\ &\leq s\left[f_0(0)\int_0^d \chi(y)dy - f_0(sA)\int_d^A [1-\chi(y)]dy\right] \\ &= s f_0(sA)\left[\frac{f_0(0)}{f_0(sA)}\int_0^d \chi(y)dy - \int_d^A [1-\chi(y)]dy\right], \end{aligned}$$

and, applying (12) and since $\lim_{s\to 0} f_0(0)/f_0(sA) = 1$, we deduce that this expression is negative for small enough s. Taking now $s = B_\chi^-(\epsilon)$ with ϵ close enough to $1-b$, we have

$$B_\chi^-(\epsilon) \leq \frac{1}{d} F_0^{-1}[1 - b/(1-\epsilon)], \qquad \text{for all} \quad d < D.$$

Therefore $BR^-(\chi) \leq a/d$ for all $d < D$, what implies $BR^-(\chi) \leq a/D$.

Proof of Theorem 4: We will apply in this case L'Hôpital Rule to obtain the implosion breakdown rate. Since $h_{\chi_a}(s) = 1 - F_0(as)$, the derivative is equal to $h'_{\chi_a}(s) = -a f_0(as)$. On the other hand, the derivative of h_χ is equal to $h'_\chi(s) = -\int_0^\infty \chi'(y) y f_0(sy)dy$ for s close enough to zero, since by the assumption on χ', we can derive under the integral sign (see Burrill (1972), p. 119). It follows that

$$\begin{aligned} \lim_{\epsilon\to 1-b} \frac{h_\chi^{-1}[b/(1-\epsilon)]}{h_{\chi_a}^{-1}[b/(1-\epsilon)]} &= \lim_{t\to 1} \frac{h_\chi^{-1}(t)}{h_{\chi_a}^{-1}(t)} = \lim_{t\to 1} \frac{h'_{\chi_a}[h_{\chi_a}^{-1}(t)]}{h'_\chi[h_\chi^{-1}(t)]} \\ &= \lim_{t\to 1} \frac{a f_0[a h_{\chi_a}^{-1}(t)]}{\int_0^\infty \chi'(y) y f_0[h_\chi^{-1}(t) y]dy} = a\left[\int_0^\infty \chi'(y) y dy\right]^{-1}, \end{aligned}$$

where we have applied the Dominated Convergence Theorem, assuming that $0 < \int_0^\infty \chi'(y)ydy < \infty$, to obtain the last equality.

The following lemmas are needed to prove some of the remaining main results. We use them to relate the explosion bias curves of the SHORTH, $B^+_{SH}(\epsilon)$, and the MAD, $B^+_M(\epsilon)$, to the bias of the median, $B_0(\epsilon)$.

Lemma 1. *Let $c = 1/F_0^{-1}(0.75)$. Suppose that F_0 satisfies Assumption 1. It holds that*

$$\lim_{\epsilon \to 0.5} \frac{B^+_{SH}(\epsilon)}{B_0(\epsilon)} = c.$$

Proof: By applying equation (5), we obtain the following expression for $B^+_{SH}(\epsilon)$:

$$B^+_{SH}(\epsilon) = cF_0^{-1}\left[\frac{3-2\epsilon}{4(1-\epsilon)}\right].$$

Therefore,

$$\lim_{\epsilon \to 0.5} \frac{B^+_{SH}(\epsilon)}{B_0(\epsilon)} = c \lim_{\epsilon \to 0.5} \frac{F_0^{-1}[(3-2\epsilon)/(4(1-\epsilon))]}{F_0^{-1}[1/(2(1-\epsilon))]}.$$

It is enough to show that the limit is one. Define $G_1(\epsilon) = F_0^{-1}[(3-2\epsilon)/(4(1-\epsilon))]$ and $G_2(\epsilon) = F_0^{-1}[1/(2(1-\epsilon))]$. We have, by the Mean Value Theorem and given that f_0 is unimodal, that

$$0 \leq \frac{G_1(\epsilon) - G_2(\epsilon)}{G_2(\epsilon)} \leq \frac{1-2\epsilon}{4(1-\epsilon)G_2(\epsilon)f_0[G_1(\epsilon)]} = \frac{1 - F_0[G_1(\epsilon)]}{G_2(\epsilon)f_0[G_1(\epsilon)]},$$

and, applying assumption 1(b), this last expression tends to zero as ϵ goes to 0.5. Therefore, $\lim_{\epsilon \to b} G_1(\epsilon)/G_2(\epsilon) = 1$.

Lemma 2. *Let $c = 1/F_0^{-1}(0.75)$. Suppose F_0 satisfies Assumption 3. It holds that*

$$\lim_{\epsilon \to 0.5} \frac{B^+_M(\epsilon)}{B_0(\epsilon)} = 2c.$$

Proof: Let $a = F_0^{-1}(0.75)$. Following Martin and Zamar (1993b), Lemma 2, $B^+_M(\epsilon)$ must verify

$$E_{F_0}\chi_a\left(\frac{X - B_0(\epsilon)}{B^+_M(\epsilon)}\right) = \frac{1-2\epsilon}{2(1-\epsilon)}, \qquad \text{where} \quad \chi_a(y) = \begin{cases} 0 & \text{if } |y| < a \\ 1 & \text{if } |y| \geq a \end{cases}. \tag{13}$$

Given that

$$E_{F_0}\chi_a\left(\frac{X - B_0(\epsilon)}{B_M^+(\epsilon)}\right) = F_0[B_0(\epsilon) - aB_M^+(\epsilon)] - F_0[B_0(\epsilon) + aB_M^+(\epsilon)] + 1,$$

It follows from (13) by reordering terms that

$$F_0[B_0(\epsilon) + aB_M^+(\epsilon)] - F_0[B_0(\epsilon) - aB_M^+(\epsilon)] = \frac{1}{2(1-\epsilon)}. \tag{14}$$

Define $H_1(\epsilon) = 2cB_0(\epsilon)$ and $H_2(\epsilon) = c[B_0(\epsilon) + B_{SH}^+(\epsilon)/c]$. We will prove that

$$H_1(\epsilon) < B_M^+(\epsilon) < H_2(\epsilon). \tag{15}$$

To demonstrate the first inequality, we substitute $H_1(\epsilon)$ by $B_M^+(\epsilon)$ in equation (14):

$$F_0[3B_0(\epsilon)] - 1 + F_0[B_0(\epsilon)] = 1/[2(1-\epsilon)] + F_0[3B_0(\epsilon)] - 1 < 1/[2(1-\epsilon)].$$

Hence, since the function $F_0[B_0(\epsilon) + ax] - F_0[B_0(\epsilon) - ax]$ is strictly increasing in x, we obtain the required inequality.

By replacing $H_2(\epsilon)$ by $B_M^+(\epsilon)$ in equation (14) we obtain

$$F_0\left[2B_0(\epsilon) + F_0^{-1}\left(\frac{3-2\epsilon}{4(1-\epsilon)}\right)\right] - \left[1 - \frac{3-2\epsilon}{4(1-\epsilon)}\right] > 2\frac{3-2\epsilon}{4(1-\epsilon)} - 1 = \frac{1}{2(1-\epsilon)}$$

and, therefore $B_M^+(\epsilon) < H_2(\epsilon)$.

From (15) and applying Lemma 1, we have that

$$2c = \lim_{\epsilon \to 0.5} \frac{H_1(\epsilon)}{B_0(\epsilon)} \leq \lim_{\epsilon \to 0.5} \frac{B_M^+(\epsilon)}{B_0(\epsilon)} \leq \lim_{\epsilon \to 0.5} \frac{H_2(\epsilon)}{B_0(\epsilon)} = 2c.$$

Proof of Theorem 5: Obvious from Lemmas 1 and 2.

Proof of Theorem 6: (a) We first find the explosion breakdown rate of S_n. In Rousseeuw and Croux (1993), Theorem 4, it is shown that

$$\Phi[B_{SH}^+(\epsilon)/c + B_S^+(\epsilon)/k] - \Phi[B_{SH}^+(\epsilon)/c - B_S^+(\epsilon)/k] = \frac{1}{2(1-\epsilon)}.$$

Defining $H_1(\epsilon) = 2kB_0(\epsilon)$ and $H_2(\epsilon) = 2kB_{SH}^+(\epsilon)/c]$ and following along the lines of the proof of Lemma 2, it is easy to show that

$$H_1(\epsilon) < B_S^+(\epsilon) < H_2(\epsilon).$$

If we divide by $B^+_{SH}(\epsilon)$, take limits when ϵ goes to 0.5 and apply Lemma 1, we obtain $BR^+(S) = 2k/c$.

Now, we are going to compute the value for the explosion breakdown rate of Q_n. From Theorem 7 of Rousseeuw and Croux (1993), we have that

$$B^+_Q(\epsilon) = d2^{1/2}\Phi^{-1}\left[\frac{5-8\epsilon+4\epsilon^2}{8(1-\epsilon)^2}\right].$$

Define $G_1(\epsilon) = (1/c)B^+_{SH}(\epsilon)$ and $G_2(\epsilon) = B^+_Q(\epsilon)/(d2^{1/2})$, and repeat the lines of the proof of Lemma 1. It follows that $\lim_{\epsilon\to 0.5} G_1(\epsilon)/G_2(\epsilon) = 1$. Therefore $BR^+(Q) = 2^{1/2}d/c$.

(b) To find $BR^-(S)$, we use Theorem 4 of Rousseeuw and Croux (1993). The implosion maximum asymptotic bias of S_n must satisfy

$$\Phi[B^-_{SH}(\epsilon)/c + B^-_S(\epsilon)/k] - \Phi[B^-_{SH}(\epsilon)/c - B^-_S(\epsilon)/k] = [(1-2\epsilon)]/[2(1-\epsilon)].$$

Applying the Mean Value Theorem to the left side yields that for each $\epsilon \in (0, 0.5)$ there exists $\alpha(\epsilon) \in (B^-_{SH}(\epsilon)/c - B^-_S(\epsilon)/k, B^-_{SH}(\epsilon)/c + B^-_S(\epsilon)/k)$ such that

$$2\varphi[\alpha(\epsilon)]B^-_S(\epsilon)/k = [(1-2\epsilon)]/[2(1-\epsilon)],$$

where φ is the density function corresponding to Φ. Therefore,

$$B^-_S(\epsilon) = \frac{k(1-2\epsilon)}{4(1-\epsilon)\varphi[\alpha(\epsilon)]}.$$

From this expression, observing that $\lim_{\epsilon\to 0.5}\alpha(\epsilon) = 0$ and applying L'Hôpital Rule, we deduce $BR^-(S) = k/c$.

Finally, to compute $BR^-(Q)$ we apply Theorem 7 in Rousseeuw and Croux (1993). $B^-_Q(\epsilon)$ satisfies

$$(1-\epsilon)^2\Phi[B^-_Q(\epsilon)/(2^{1/2}d)] + 2\epsilon(1-\epsilon)\Phi[B^-_Q(\epsilon)] + \epsilon^2 = 5/8. \tag{16}$$

Using the Mean Value Theorem again, it follows that there exist $\alpha = \alpha(\epsilon) \in (0, B^-_Q(\epsilon)/(2^{1/2}d)$ and $\beta = \beta(\epsilon) \in (0, B^-_Q(\epsilon))$ such that

$$\begin{aligned}\Phi[B^-_Q(\epsilon)/(2^{1/2}d)] &= 1/2 + \varphi(\alpha)B^-_Q(\epsilon)/(2^{1/2}d) \quad \text{and}\\ \Phi[B^-_Q(\epsilon)] &= 1/2 + \varphi(\beta)B^-_Q(\epsilon).\end{aligned}$$

Plugging these expressions in equation (16), we have

$$(1-\epsilon)^2[1/2 + \varphi(\alpha)B^-_Q(\epsilon)/(2^{1/2}d)] + 2\epsilon(1-\epsilon)[1/2 + \varphi(\beta)B^-_Q(\epsilon)] + \epsilon^2 = 5/8.$$

Some tedious calculations yield

$$B_Q^-(\epsilon) = \frac{2^{1/2}d(1-4\epsilon^2)}{8(1-\epsilon)^2\varphi(\alpha) + 2^{1/2}16d\epsilon(1-\epsilon)\varphi(\beta)}.$$

Notice that $\lim_{\epsilon\to 0.5}\alpha(\epsilon) = \lim_{\epsilon\to 0.5}\beta(\epsilon) = 0$. From an application of L'Hôpital Rule and some calculations it follows that $RB^-(Q) = c^{-1}[1+2^{-3/2}d^{-1}]^{-1}$.

In Lemma 3 we compute the breakdown rate for a special location M–estimate. This lemma will be used in the proof of Theorem 7.

Lemma 3. *Let* $c = 1/F_0^{-1}(0.75)$. *Define*

$$\psi_A(y) = \begin{cases} 0 & \textit{if } |y| < A \\ sgn(y) & \textit{if } |y| \geq A \end{cases}.$$

The breakdown rate of the location M–estimate defined by ψ_A *is*

$$BR(\psi_A) = 1 + cD^+(\epsilon)A.$$

Proof: Let $B_A(\epsilon)$ be the maximum bias of the M–estimate defined by ψ_A. Equation (9) says that $B_\psi(\epsilon) = g_\psi^{-1}[\epsilon/(1-\epsilon), D^+(\epsilon)]$, where

$$\begin{aligned} g_A(t,s) &= s\int_0^\infty \psi_A(y)(f_0(sy-t) - f_0(sy+t))dy \\ &= F_0(As+x) - F_0(As-x). \end{aligned}$$

By taking $t = B_A(\epsilon)$ and $s = D^+(\epsilon)$ in the last equation, it follows that

$$\epsilon/(1-\epsilon) = F_0[AD^+(\epsilon) + B_A(\epsilon)] - F_0[AD^+(\epsilon) - B_A(\epsilon)].$$

Define $H_1(\epsilon) = AD^+(\epsilon) + F_0^{-1}[\epsilon/(1-\epsilon)]$ and $H_2(\epsilon) = AD^+(\epsilon) + B_0(\epsilon)$. The same arguments as in the proof of Lemma 2 show that $H_1(\epsilon) < B_A(\epsilon) < H_2(\epsilon)$. If we divide now by $B_0(\epsilon)$ in this inequalities, take limits as ϵ goes to 0.5 and apply Lemma 1, we obtain

$$cBR^+(D)A + \lim_{\epsilon\to 0.5}\frac{F_0^{-1}[\epsilon/(1-\epsilon)]}{B_0(\epsilon)} \leq BR(\psi_k) \leq cBR^+(D) + 1.$$

Therefore, it is enough to show

$$\lim_{\epsilon\to 0.5}\frac{F_0^{-1}[\epsilon/(1-\epsilon)]}{B_0(\epsilon)} = 1.$$

But this is a consequence of Assumption 1. We must only define $G_1(\epsilon) = F_0^{-1}[1/(2(1-\epsilon))]$ and $G_2(\epsilon) = F_0^{-1}[\epsilon/(1-\epsilon)]$ and repeat the same argument as in the proof of Lemma 1.

Proof of Theorem 6: (a) First, we show that $1 + cBR^+(D)A$ is an upper bound for $BR(\psi)$. Let ψ_A be the function defined in Lemma 3. Then, it is clear that $\psi_A(x) \leq \psi(x)$ for every $x \geq 0$. This implies that $g_A(t,s) \leq g_\psi(t,s)$ for $s > 0$ and $0 < t < 1$. Therefore, due to equation (9), we have $B_\psi(\epsilon) \leq B_A(\epsilon)$. By applying Lemma 3, it follows that $BR(\psi) \leq BR(\psi_A) = 1 + cBR^+(D)A$.

Now, we will show that $cBR^+(D)A$ is a lower bound for $BR(\psi)$. Define for $0 < r < A$ the following ancillary function

$$\psi_r(y) = \begin{cases} \mathrm{sgn}(y) & \text{if } |y| > r \\ \psi(r)\mathrm{sgn}(y) & \text{if } |y| \leq r \end{cases}.$$

After some manipulations we have that,

$$g_r(t,s) = [1-\psi(r)][F_0(rs+t) - F_0(rs-t)] + \psi(r)[2F_0(t) - 1],$$

for $0 < t < 1$ and $s > 0$. By substituting $B_r(\epsilon)$ for t and $D^+(\epsilon)$ for s in the last equation we obtain for $0 < \epsilon < 0.5$

$$\begin{aligned} \epsilon/(1-\epsilon) &= [1-\psi(r)][F_0(rD^+(\epsilon) + B_r(\epsilon)) - F_0(rD^+(\epsilon) - B_r(\epsilon))] \\ &+ \psi(r)[2F_0(B_r(\epsilon)) - 1]. \end{aligned}$$

Now, we take limits as ϵ goes to 0.5 in both sides of this equation. The left hand side converges to one and the same must happen to the right member. This implies in particular that $\lim_{\epsilon \to 0.5}[rD^+(\epsilon) - B_r(\epsilon)] = -\infty$. So, there exists ϵ_0 such that if $\epsilon_0 < \epsilon < 0.5$, then $B_r(\epsilon) > rD^+(\epsilon)$. It follows from Lemma 1 that

$$BR(\psi_r) \geq cBR^+(D)r. \tag{17}$$

If $0 < r < A$, we have that $\psi(y) \leq \psi_r(y)$ for $y \geq 0$. Hence $B_\psi(\epsilon) \geq B_r(\epsilon)$, following the lines of the first part of the proof. Therefore, for every $0 < r < A$, $BR(\psi) \geq BR(\psi_A) \geq cBR^+(D)r$. Now, if $r \to A$, it follows that $BR(\psi) \geq cBR^+(D)A$.

(b) Since $\psi(y) < 1$ for each y, it holds that $\psi(y) \leq \psi_r(y)$ for $y \geq 0$ and $r \geq 0$. Therefore $B_\psi(\epsilon) \geq B_r(\epsilon)$ for $r > 0$. It follows applying inequality (17) that $BR(\psi) \geq BR(\psi_r) \geq cBR^+(D)r$ for $r > 0$. This yields the desired result.

Acknowledgements

José R. Berrendero and Juan Romo were partially supported by PB93–0232 grant from DGCYT (Spain) and Rubén Zamar was partially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by PB92–0247 grant from DGICYT (Spain).

References

Burrill, C.W. (1972). *Measure, Integration and Probability.* McGraw-Hill, New York.

Croux C. (1994). Efficient high-breakdown M-estimators of scale. *Statist. Probab. Lett.* **19**, 371–379.

Hampel, F.R. (1968). Contributions to the theory of robust estimation. Ph.D. thesis. University of California, Berkeley.

Hampel, F.R. (1974). The influence curve and its role in robust estimation. *J. Am. Statist. Assoc.* **69**, 383-393.

Hössjer, O. (1992). On the optimality of S-estimators. *Statist. Probab. Lett.* **14**, 413–419.

Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35**, 73–101.

Huber, P.J. (1981). *Robust Statistics.* Wiley, New York.

Martin, R.D. and Zamar, R.H. (1989). Asymptotically Min-Max bias robust M–estimates of scale for positive random variables. *J. Am. Statist. Assoc.* **17**, 494–501.

Martin, R.D. and Zamar, R.H. (1993a). Efficiency-constrained bias-robust estimation of location. *Ann. Statist.* **21**, 338–354.

Martin, R.D. and Zamar, R.H. (1993b). Bias robust estimation of scale. *Ann. Statist.* **21**, 991–1017.

Martin, R.D., Yohai, V.J. and Zamar, R.H. (1989). Min-Max bias robust regression. *Ann. Statist.* **17**, 1608–1630.

Mazzi, S.V.T. (1992). A new measure of quantitative robustness. M. Sc. Thesis. University of British Columbia.

Rousseeuw, P.J. and Croux, C. (1993). Alternatives to the Median Absolute Deviation. *J. Am. Statist. Assoc.* **88**, 1273–1283.

Rousseeuw, P.J. and Croux, C. (1994). The bias of k-step M-estimators. *Statist Probab. Lett.* **20**, 411–420.

Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust regression and outlier detection.* Wiley, New York.

Rousseeuw, P.J. and Yohai, V.J. (1984). Robust regression by means of S-estimators. In *Robust and Nonlinear Time Series Analysis*, eds. J. Franke, W. Härdle and R.D. Martin. Lecture Notes in Statistics 26, Springer, New-York, 256–272.

Yohai, V.J. and Zamar, R.H. (1988). High breakdown point estimates of regression by means of the minimization of an efficient scale. *J. Amer. Statist. Assoc.* **83**, 406–413.

Yohai, V. and Zamar, R. (1992). Optimally bounding a generalized gross error sensitivity of unbounded influence M–estimates of regression. Working Paper, 92-44. Universidad Carlos III de Madrid, Spain.

ADDRESS FOR CORRESPONDENCE:
JUAN ROMO
DEPARTAMENTO DE ESTADÍSTICA Y ECONOMETRÍA
UNIVERSIDAD CARLOS III DE MADRID
C/ MADRID 126
28903 GETAFE
MADRID SPAIN

Function	Tuning constant	γ^*
$\chi_H(y) = \min(y^2/c^2, 1)$	1.041	1.23
$\chi_T(y) = \min(3y^2/c^2 - 3y^4/c^4 + y^6/c^6, 1)$	1.547	1.28
$\chi_L(y) = \min(\lvert y/c \rvert, 1)$	1.470	1.39

Table 1. Score functions, tuning constants and gross error sensitivities for three well-known scale M–estimates.

Function	k	$RBR^{+}(\chi, \chi_H)$	$RBR^{-}(\chi, \chi_H)$	γ^*	EFF
χ_{CR}	1	∞	0.73	2.69	76.7
χ_L	1	∞	0.94	1.39	61.5
χ_C	2	1.70	0.72	1.59	52.2
χ_W	2	1.27	0.96	2.66	54.9
χ_T	2	1.16	0.99	1.28	53.9
χ_H	2	1	1	1.23	50.6
χ_4	4	0	1.03	1.19	43.9

Table 2. Relative explosion and implosion breakdown rate for several scale M–estimates taking Huber M–estimate as the baseline. Gross error sensitivity and asymptotic efficiency are also displayed.

Estimate	BR^+	BR^-	γ^*	EFF
$SHORTH$	1	1	1.16	36.74
MAD	2	1	1.16	36.74
S	1.60	0.80	1.62	58.23
Q	2.12	0.58	2.06	82.27

Table 3. Explosion and implosion breakdown rate, gross error sensitivity and asymptotic efficiency of four robust dispersion estimates.

A	$SHORTH$	MAD
0.5	1.37	1.99
1	1.99	3.48
1.25	2.36	4.22
1.5	2.73	4.97
1.75	3.10	5.71

Table 4. Middle point between the lower and the upper bounds for several location M–estimates.

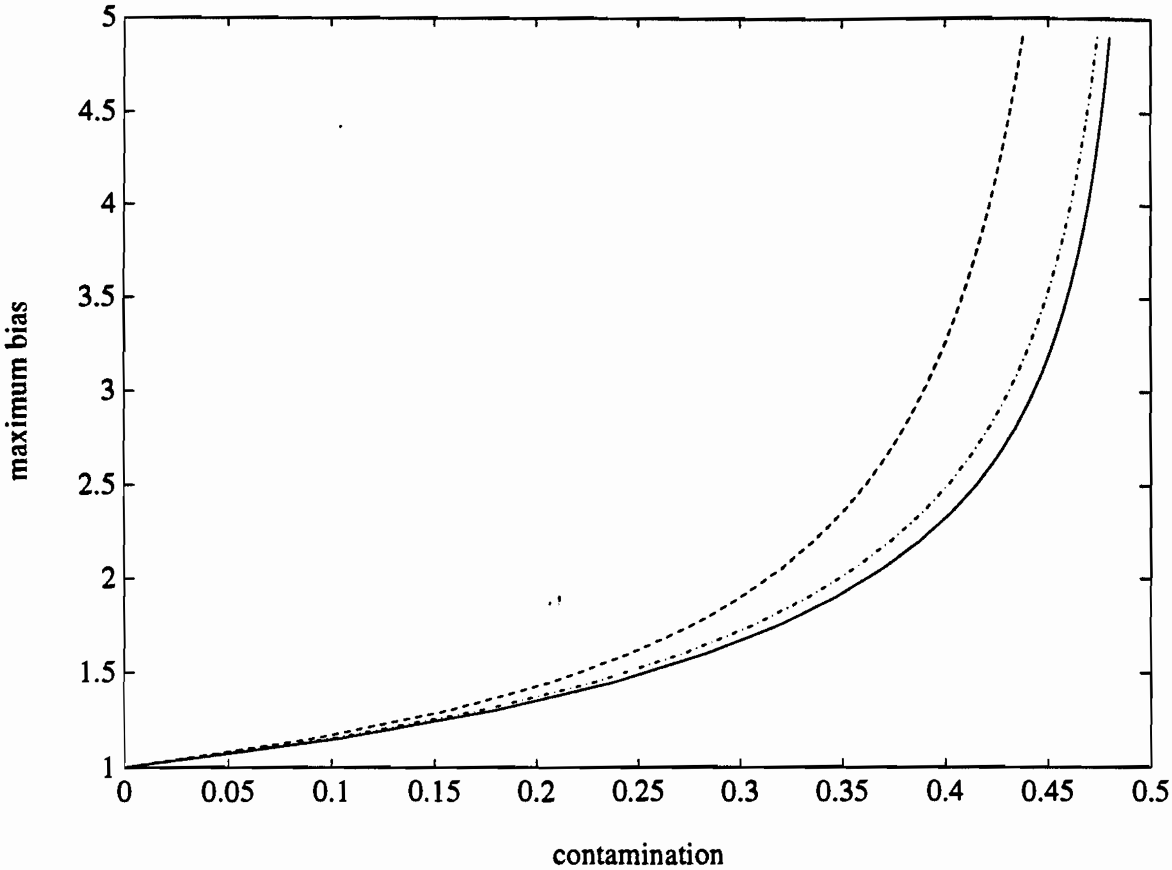

Figure 1. Explosion maximum asymptotic bias curves for $\widehat{\sigma}_L$ (dashed line), $\widehat{\sigma}_T$ (dotted-dashed line), and $\widehat{\sigma}_H$ (solid line).

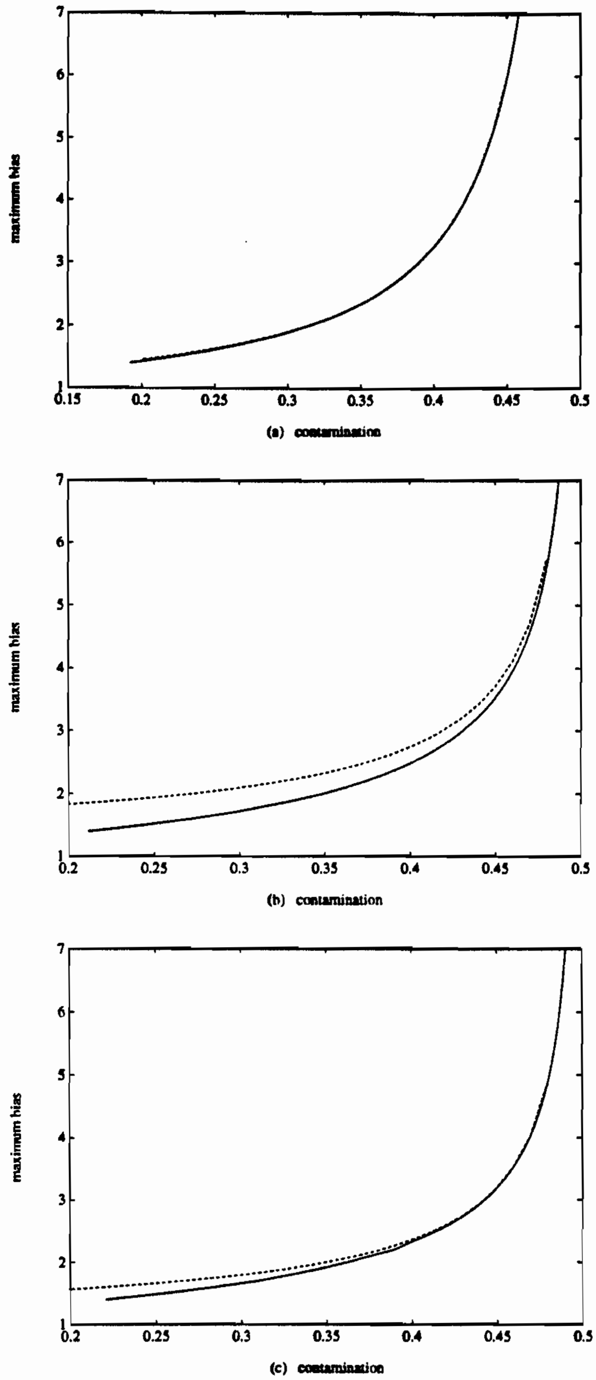

Figure 2. Approximation (dashed line) to the maximum bias curves (solid line) for $\hat{\sigma}_L$ (a), $\hat{\sigma}_T$ (b), and $\hat{\sigma}_H$ (c), obtained from Theorem 1.

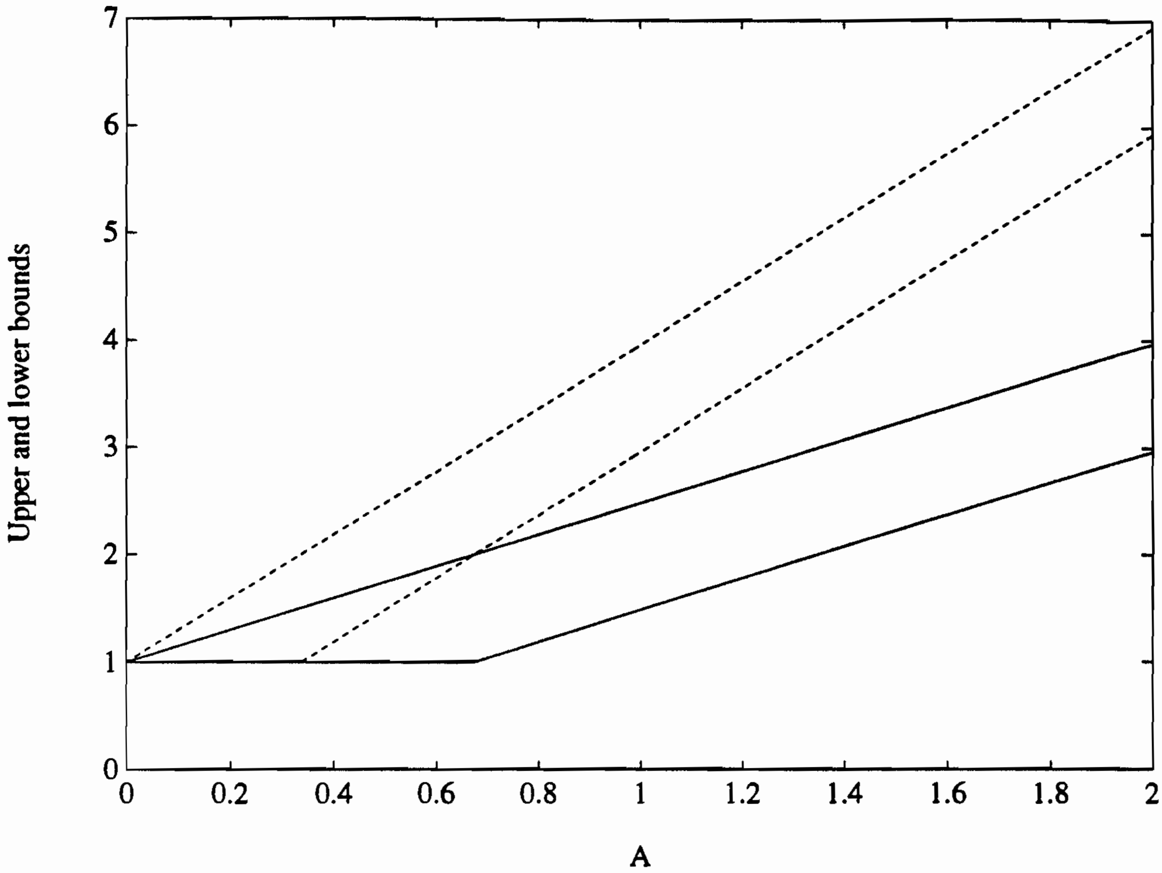

Figure 3. Upper and lower bounds for $BR(\psi)$ when the dispersion estimates are the MAD (dashed line) and the SHORTH (solid line).